

NAACL HLT 2019

**The Workshop on Computational Approaches to  
Subjectivity, Sentiment and Social Media Analysis**

**Proceedings of the Tenth Workshop**

June 6, 2019  
Minneapolis, USA

©2019 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)  
209 N. Eighth Street  
Stroudsburg, PA 18360  
USA  
Tel: +1-570-476-8006  
Fax: +1-570-476-0860  
[acl@aclweb.org](mailto:acl@aclweb.org)

ISBN 978-1-950737-12-3

## Introduction

Two thousand five hundred years ago, Democritus said: “Nothing exists except atoms and empty space. Everything else is opinion.” Today, this saying is omnipresent, as the state of our social interpretation of reality has been denominated by the “post-truth society”. Research in automatic Subjectivity and Sentiment Analysis (SSA), as subtasks of Affective Computing and Natural Language Processing (NLP), has flourished in the past years. The growth in interest in these tasks was motivated by the birth and rapid expansion of the Social Web that made it possible for people all over the world to share, comment or consult content on any given topic. In this context, opinions, sentiments and emotions expressed in Social Media texts have been shown to have a high influence on social and economic behaviour worldwide.

SSA systems are highly relevant to many real-world applications (e.g. marketing, eGovernance, business intelligence, social analysis) and also to many tasks in Natural Language Processing (NLP), e.g. information extraction, question answering, textual entailment, to name just a few. The importance of this field has been proven by the high number of approaches proposed in research in the past decade, as well as by the interest that it raised from other disciplines (Economics, Sociology, Psychology, Marketing, Crisis Management, Digital Humanities and Behavioral Studies) and the applications that were created using its technology. Next to the growth in the diversity of applications, task definitions changed towards more complex challenges: Subjectivity, polarity recognition and opinion mining has been enriched with fine-grained aspect and target level predictions. Polarity as a concept is complemented by emotion models as defined from psychological research. In spite of the growing body of research in the area in the past years, dealing with affective phenomena in text has proven to be a complex, interdisciplinary problem that remains far from being solved. Its challenges include the need to address the issue from different perspectives and at different levels, depending on the characteristics of the textual genre, the language(s) treated and the final application for which the analysis is done.

The aim of the 10th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2019) was to continue the line of the previous editions, bringing together researchers in Computational Linguistics working on Subjectivity and Sentiment Analysis and researchers working on interdisciplinary aspects of affect computation from text.

For this year’s workshop, we accepted 12/19 papers (63 % acceptance rate). The number of submissions has been comparably low in contrast to previous years, given that WASSA 2018 just took place a little more than half a year before WASSA 2019. Nevertheless and given the high quality of submissions, we have an interesting program consisting of 11 oral presentations in 3 sessions. An invited talk by Sven Buechel on "Modeling Empathy and Distress in Reaction to News Stories" complements this year’s paper presentations.

Accepted papers deal with a variety of topics, like stance detection, topic and aspect level sentiment analysis, social network analysis, humor detection, and negations and intensifiers in emotion and sentiment analysis, as well as applications of such systems to real-world problems. For the first time, we asked the reviewers for recommendations for a best paper award. Based on these recommendations, the organizers voted secretly on the best paper selection. The result is that the paper “Stance Detection in Code-Mixed Hindi-English Social Media Data using Multi-Task Learning” by Sushmitha Reddy Sane, Suraj Tripathi, Koushik Reddy Sane and Radhika Mamidi wins this year’s best paper award.

We would like to thank the NAACL 2019 Organizers and Workshop Chairs for their help and support at the different stages of the workshop organization process. We are also especially grateful to the Program Committee members for the time and effort they spent in assessing the papers.

Alexandra Balahur, Roman Klinger, Véronique Hoste, Carlo Strapparava and Orphée De Clercq  
WASSA 2019 Chairs

# **Keynote talk: Modeling Empathy and Distress in Reaction to News Stories**

**Sven Buechel**

JULIE Lab, Jena University, Germany

## **Abstract**

Computational detection and understanding of empathy is an important factor in advancing human-computer interaction. Yet to date, text-based empathy prediction has the following major limitations: It underestimates the psychological complexity of the phenomenon, adheres to a weak notion of ground truth where empathic states are ascribed by third parties, and lacks a shared corpus. In contrast, this talk describes the first publicly available gold standard for empathy prediction. It is constructed using a novel annotation methodology which reliably captures empathy assessments by the writer of a statement using multi-item scales. This is also the first computational work distinguishing between multiple forms of empathy, empathic concern, and personal distress, as recognized throughout psychology.

## Organizers

### Organizers:

Alexandra Balahur – European Commission Joint Research Centre  
Roman Klinger – University of Stuttgart, Germany  
Véronique Hoste – Ghent University, Belgium  
Carlo Strapparava – Fondazione Bruno Kessler, Italy  
Orphée de Clercq – Ghent University, Belgium

### Program Committee:

Jeremy Barnes – University of Oslo  
Sabine Bergler – Concordia University  
Cristina Bosco – Dipartimento di Informatica - Università di Torino  
Laura Ana Maria Bostan – IMS, University of Stuttgart  
Sven Buechel – Friedrich-Schiller-Universität Jena  
Erik Cambria – Nanyang Technological University  
Luna De Bruyne – LT3, Language and Translation Technology Team, Ghent University  
Lingjia Deng – Bloomberg L.P.  
Antske Fokkens – VU Amsterdam  
Michael Gamon – Microsoft Research  
Lorenzo Gatti – Human Media Interaction, University of Twente  
Matthias Hartung – Semalytix GmbH  
Carlos A. Iglesias – Universidad Politécnica de Madrid  
Gilles Jacobs – LT3, Language and Translation Technology Team, Ghent University  
Aditya Joshi – CSIRO  
Manfred Klenner – Computational Linguistics, University of Zurich  
Els Lefever – LT3, Language and Translation Technology Team, Ghent University  
Gerard Lynch – University College Dublin  
Isa Maks – VU University of Amsterdam  
Edison Marrese-Taylor – The University of Tokyo  
Saif Mohammad – National Research Council Canada  
Karo Moilanen – Dpt of Computer Science, University of Oxford  
Malvina Nissim – University of Groningen  
Sean Papay – University of Stuttgart  
Viviana Patti – Dipartimento di Informatica, University of Turin  
Jose Manuel Perea-Ortega – University of Extremadura  
Daniel Preotiu-Pietro – Bloomberg  
Paolo Rosso – Universitat Politècnica de València  
Alon Rozental – Amobee inc.  
Mohammad Salameh – Diffbot  
Josef Steinberger – University of West Bohemia  
Mike Thelwall – University of Wolverhampton  
Dan Tufiş – Research Institute for Artificial Intelligence, Romanian Academy  
Michael Wiegand – Saarland University  
Michael Wojatzki – Language Technology Lab, University of Duisburg-Essen  
Taras Zagibalov – SmallStep LLC



## Table of Contents

<i>Stance Detection in Code-Mixed Hindi-English Social Media Data using Multi-Task Learning</i> Sushmitha Reddy Sane, Suraj Tripathi, Koushik Reddy Sane and Radhika Mamidi .....	1
<i>A Soft Label Strategy for Target-Level Sentiment Classification</i> Da Yin, Xiao Liu, Xiuyu Wu and Baobao Chang .....	6
<i>Online abuse detection: the value of preprocessing and neural attention models</i> Dhruv Kumar, Robin Cohen and Lukasz Golab .....	16
<i>Exploring Fine-Tuned Embeddings that Model Intensifiers for Emotion Analysis</i> Laura Ana Maria Bostan and Roman Klinger .....	25
<i>Enhancing the Measurement of Social Effects by Capturing Morality</i> Rezvaneh Rezapour, Saumil H. Shah and Jana Diesner .....	35
<i>Using Structured Representation and Data: A Hybrid Model for Negation and Sentiment in Customer Service Conversations</i> Amita Misra, Mansurul Bhuiyan, Jalal Mahmud and Saurabh Tripathy .....	46
<i>Deep Learning Techniques for Humor Detection in Hindi-English Code-Mixed Tweets</i> Sushmitha Reddy Sane, Suraj Tripathi, Koushik Reddy Sane and Radhika Mamidi .....	57
<i>How do we feel when a robot dies? Emotions expressed on Twitter before and after hitchBOT's destruction</i> Kathleen C. Fraser, Frauke Zeller, David Harris Smith, Saif Mohammad and Frank Rudzicz ....	62
<i>"When Numbers Matter!": Detecting Sarcasm in Numerical Portions of Text</i> Abhijeet Dubey, Lakshya Kumar, Arpan Somani, Aditya Joshi and Pushpak Bhattacharyya ....	72
<i>Cross-lingual Subjectivity Detection for Resource Lean Languages</i> Ida Amini, Samane Karimi and Azadeh Shakery .....	81
<i>Analyzing Incorporation of Emotion in Emoji Prediction</i> Shirley Anugrah Hayati and Aldrian Obaja Muis .....	91





# Conference Program

**Thursday, June 6, 2019**

**9:00–9:15**     *Opening of the WASSA Workshop by Alexandra Balahur*

**9:15–10:15**     *Invited Talk by Sven Buechel: Modeling Empathy and Distress in Reaction to News Stories*

**10:15–10:45**     *Morning Coffee Break*

## **Oral Session 1**

**10:45–11:10**     *Stance Detection in Code-Mixed Hindi-English Social Media Data using Multi-Task Learning*  
Sushmitha Reddy Sane, Suraj Tripathi, Koushik Reddy Sane and Radhika Mamidi

**11:10–11:35**     *A Soft Label Strategy for Target-Level Sentiment Classification*  
Da Yin, Xiao Liu, Xiuyu Wu and Baobao Chang

**11:35–12:00**     *Online abuse detection: the value of preprocessing and neural attention models*  
Dhruv Kumar, Robin Cohen and Lukasz Golab

**12:00–12:25**     *Exploring Fine-Tuned Embeddings that Model Intensifiers for Emotion Analysis*  
Laura Ana Maria Bostan and Roman Klinger

**12:25–14:00**     *Lunch Break*

**Thursday, June 6, 2019 (continued)**

**Oral Session 2**

- 14:00–14:25 *Enhancing the Measurement of Social Effects by Capturing Morality*  
Rezvaneh Rezapour, Saumil H. Shah and Jana Diesner
- 14:25–14:50 *Using Structured Representation and Data: A Hybrid Model for Negation and Sentiment in Customer Service Conversations*  
Amita Misra, Mansurul Bhuiyan, Jalal Mahmud and Saurabh Tripathy
- 14:50–15:15 *Deep Learning Techniques for Humor Detection in Hindi-English Code-Mixed Tweets*  
Sushmitha Reddy Sane, Suraj Tripathi, Koushik Reddy Sane and Radhika Mamidi
- 15:15–15:40 *How do we feel when a robot dies? Emotions expressed on Twitter before and after hitchBOT's destruction*  
Kathleen C. Fraser, Frauke Zeller, David Harris Smith, Saif Mohammad and Frank Rudzicz

**15:40–16:10 Afternoon Coffee Break**

**Oral Session 3**

- 16:10–16:35 *"When Numbers Matter!": Detecting Sarcasm in Numerical Portions of Text*  
Abhijeet Dubey, Lakshya Kumar, Arpan Somani, Aditya Joshi and Pushpak Bhat-tacharyya
- 16:35–17:00 *Cross-lingual Subjectivity Detection for Resource Lean Languages*  
Ida Amini, Samane Karimi and Azadeh Shakery
- 17:00–17:25 *Analyzing Incorporation of Emotion in Emoji Prediction*  
Shirley Anugrah Hayati and Aldrian Obaja Muis

**17:25–18:00 Closing Remarks**